

# 제한된 기상변수와 Polynomial Networks Approach를 이용한 일 증발접시 증발량의 모형화

## Modeling of Daily Pan Evaporation using the Limited Climatic Variables and Polynomial Networks Approach

김성원\*

### Abstract

Group method of data handling neural networks model (GMDH-NNM) is used to estimate daily pan evaporation (PE) using limited climatic variables such as max temperature ( $T_{max}$ ), min temperature ( $T_{min}$ ), mean wind speed ( $W_{mean}$ ), mean relative humidity ( $RH_{mean}$ ) and sunshine duration (SD). And, for the performances of GMDH-NNM, it is composed of training and test performances, respectively. The training and test performances are carried out using daily time series data, respectively. From this research, we evaluate the impact of GMDH-NNM for the modeling of the nonlinear time series data. We should, thus, construct the credible data of the daily PE data using GMDH-NNM, and can suggest the methodology for the irrigation and drainage networks system. Furthermore, this research represents that the strong nonlinear relationship such as pan evaporation modeling can be generalized using GMDH-NNM.

*Key words* : Polynomial Networks Approach, GMDH-NNM, Limited Climatic Variables, Pan Evaporation

### 1. 서론

증발은 수문학, 농업경제학, 관개운영, 하천유출량 예측 및 호소의 물수지의 모형화와 같은 많은 분야에서 이용되어 지는 근본적인 인자이다 (Terzi and Keskin, 2005; Lenters et al., 2005) Class A형 증발접시는 증발량 측정을 위하여 가장 광범위하게 사용되는 도구 중에 하나이다. 그러나 이러한 도구를 사용하는 데는 많은 문제점을 가지고 있다. 수체의 파편, 증발접시 내부와 주변에서의 동물들의 활동, 증발접시의 크기, 증발접시를 만드는데 사용한 재료, 증발접시의 노출상태 및 강한 바람 등은 증발량의 측정에 있어서의 오차를 나타낼 수 있다 (Vining, 2003). 증발접시 증발량의 관측의 경우 선진국에서는 특히 시·공간적으로 제한적이다. 증발접시 증발량의 관측이 부족한 지역에서는 지역적인 광범위한 조건에 강하게 의존하는 특정 기상학적 자료를 이용하여 이론적 및 경험적인 추정치로부터 증발접시 증발량을 예측한다 (Rahimi Khoob, 2009; Sabziparvar et al., 2009).

본 연구의 목적은 제한된 기상학적 변수와 polynomial networks approach를 이용하여 일 증발접시 증발량의 모형화에 있다. polynomial networks approach를 이용한 신경망모형의 하나인 group method of data handling neural networks model (GMDH-NNM)의 수행평가를 위하여 모형화의 과정은 훈련 및 테스트과정으로 구성되었다. 본 연구에서는 비선형 수문학적 시계열자료의 모형화를 위해서 제한된 서로 다른 기상변수를 이용하여 GMDH-NNM의 모형화에 대한 적용성을 평가하였다.

### 2. 신경망모형

#### 2.1 자료처리 군집화기법 신경망모형 (Group Method of Data Handling Neural Networks Model, GMDH-NNM)

본 연구에 적용된 자료처리 군집화기법 신경망모형 (GMDH-NNM)은 다항식의 네트워크를 가지는 강력한 구조를 나타내고 있다. GMDH-NNM은 얼마나 많은 층을 구조화 할 것인지를 결정하기 위하여 유전적인 인자와 연결에서 다항식의 항으로 나타내어 질 수 있다. GMDH-NNM의 주요 개념은 실측된 출력층 노드값에 가능한 한 가장 근접한 예측된 출력층 노드값을 계산하기 위한 함수를 설정하는 것이라고 할 수 있다. 많은 전문가들은 그동안 광범위하게 사용되어 왔으며, 널리 이해되고 있는 다항식을 사용하여 모형화 및 예측을 할 수 있는 방법론을 가진다는 것이 훨씬 편할 수 있다. GMDH-NNM을 해결하기 위한 가장 보편적인 접근법은 회귀분석을 이용하는 것이다. 회귀분석을 실시하는 첫 번째 단계는 회귀식이 유도되어야 하는 다항식의 형태를 결정하는 것이다. 두 번째 단계는 입력층 노드를 이용하여 모든 다항식항의 선형조합을 설정하는 것이다. 즉 알고리즘은 출력층 노드의 실측치와 계산치 사이에 오차제곱합을 최소화 함으로써 이러한 계수들의 합을 결정한다. 회귀식을 이용할 경우에 주요 문제는 정확하게 다항식의 항 자료를 어떻게 결정하느냐가 주요 문제가 된다. 즉 항을 얼마나 복잡하게 설정하느냐

\* 정회원, 공학박사, 수자원개발기술사, 동양대학교 철도토목학과 부교수 ; E-mail : [swkim1968@dyu.ac.kr](mailto:swkim1968@dyu.ac.kr)

나가 가장 중요하다. 예를 들면 하나의 변수에 대한 모형 다항식이 포함된다면, 다항식의 최대 정확도는 어떻게 설정하여야 하는 것이다. 그러나 GMDH-NNM은 가능한 한 모든 조합을 시도하여 보편적인 회귀식에 대한 경우보다 최적의 다항식을 형성할 수 있다 (Farlow, S.J., 1984; Yoshimura, et al., 1985; Neuroshell 2; 1993). 본 연구에서 GMDH-NNM은 Scale은 최소값 -1, 최대값 1로서 선형으로 구성하였으며, SMART GMDH-NNM 모형으로 선정하였다. 그리고 모형의 비선형성은 최대로 설정하였다.

### 3. 연구범위 및 자료

#### 3.1 연구범위

본 연구에서 기상관측소는 제주도를 포함한 우리나라 기상청의 관리하에 있는 71개의 기상관측소중에서 선정하였다. 선택된 기상관측소는 먼저 전 국토에 골고루 분포되어 있어서 그 지역을 대표할 수 있는 지점이어야 하며, 최소 30년 이상의 기상관측 자료를 보유하고 있어야 한다. 따라서 이러한 조건에 부합되는 수 많은 기상관측소 중에서 광주기상대 및 대기기상대를 선택하였다.

#### 3.2 연구범위

본 연구에 개발된 SMART GMDH-NNM을 적용하기 위하여 필요한 기상자료는 국토해양부의 국가 수자원관리종합정보 시스템(Water Management Information System, WAMIS)의 홈페이지([www.wamis.go.kr](http://www.wamis.go.kr))와 기상청의 기상청 홈페이지([www.kma.go.kr](http://www.kma.go.kr))의 과거 무료 공개자료로부터 필요한 자료를 이용하였다. 그리고 SMART GMDH-NNM을 적용하기 위하여 이용한 기상자료는 1985년 1월 1일부터 1990년 12월 31일까지의 일단위의 최대기온 ( $T_{max}$ ), 최소기온 ( $T_{min}$ ), 평균풍속 ( $W_{mean}$ ), 평균 상대습도 ( $RH_{mean}$ ), 일조시간 (SD) 및 증발점시 증발량 (PE)의 기상자료를 이용하여 훈련자료 및 테스트자료로 구성하였다.

## 4. SMART GMDH-NNM의 적용

### 4.1 훈련과정 (Training Performance)

본 연구에서 훈련과정은 광주기상대 및 대기기상대로부터 1985년 1월 1일부터 1989년 12월 31일까지의 일단위의 기상자료를 이용하였으며, 총 1,826개의 자료시계열로 구성되어 있다. 또한 SMART GMDH-NNM은 광주기상대 및 대기기상대에서 각각 5개의 모형으로 구성하였다. 모형의 구성은 입력층 노드를 구성한 기상인자에 따라서 구분되었으며, 출력층 노드는 증발점시 증발량으로만 구성되었다. 다음 Table 1은 본 연구에서 구성된 SMART GMDH-NNM의 특정기상대별 모형의 구조, 입력층 노드 및 출력층 노드를 나타내고 있다. 그리고 SMART GMDH-NNM의 훈련에 따른 선행과정으로서 입력층 노드의 각 기상자료를 표준화하여 적용하였다. 일반적으로 입력층 노드들을 표준화하는 주요한 이유 중 하나는 각 입력층 노드들이 서로 다른 단위로서 관측된 값들을 나타내고 있기 때문에, 이러한 입력층 노드들을 표준화하고 무차원단위에서 입력층 노드들을 재배치함으로써 각 변수들의 상사성의 효과를 제거하기 위한 것으로 판단된다 (Kim and Kim, 2008a, b; Kim et al., 2009). 또한 다음 Table 2는 광주기상대에서 각 모형별 훈련결과 산정된 최적 다항식 및 구성된 층을 나타내고 있다. 각 모형별 훈련결과 광주기상대 및 대기기상대에서 5개의 입력층 노드로 구성된 KWA-MOD-5 및 DAE-MOD-5의 훈련결과가 가장 양호한 것으로 나타났다.

### 4.2 테스트과정 (Test Performance)

본 연구에서 테스트과정은 광주기상대 및 대기기상대로부터 1990년 1월 1일부터 1990년 12월 31일까지의 일단위의 기상자료를 이용하였으며, 총 365개의 자료시계열로 구성되어 있다. 따라서 훈련과정 동안 산정된 최적 다항식을 이용하여 테스트과정에서 증발점시 증발량을 모형화하였다. 그 결과 훈련과정과 동일하게 광주기상대 및 대기기상대에서 5개의 입력층 노드로 구성된 KWA-MOD-5 및 DAE-MOD-5의 테스트결과가 가장 양호한 것으로 나타났다. 다음 Figure 1(a)-(c)는 광주기상대의 KWA-MOD-3, KWA-MOD-4 및 KWA-MOD-5 모형의 훈련 및 테스트과정 결과 증발점시 증발량의 관측치와 계산치의 비교결과를 나타내고 있으며, 다음 Figure 2(a)-(c)는 대기기상대의 KWA-MOD-3, KWA-MOD-4 및 KWA-MOD-5 모형의 훈련 및 테스트과정 결과 증발점시 증발량의 관측치와 계산치의 비교결과를 나타내고 있다.

## 5. 결 론

본 연구에서는 SMART GMDH-NNM과 제한된 기상자료를 이용하여 증발점시 증발량을 모형화하였다. 적용된 기상자료는 광주기상대 및 대기기상대의 기상자료인 1985년 1월 1일부터 1990년 12월 31일까지의 일단위의 최대기온 ( $T_{max}$ ), 최소기온 ( $T_{min}$ ), 평균풍속 ( $W_{mean}$ ), 평균 상대습도( $RH_{mean}$ ), 일조시간 (SD) 및 증발점시 증발량 (PE)의 기상자료를 이용하였으며, 입력층 노드는 각 기상변수의 조합으로 구성하였다. 또한 출력층 노드는 증발점시 증발량 하나만 구성되어 있다. 본 연구의 훈련 및 테스트 과정을 통하여 광주기상대와 대기기상대의 KWA-MOD-5 및 DAE-MOD-5의 모형이 가장 최적의 결과를 나타내는 것으로 분석되었다. 그러나 본 연구는 2개의 기상관측소 및 5개의 입력층 노드의 조합만으로 SMART GMDH-NNM의 모형화를 평가하였다.

즉 SMART GMDH-NNM와 제한된 기상변수만으로 증발접시 증발량의 모형화에 대한 적용성을 검토하였으며, 그에 따른 잠재성을 분석하였다. 그러나 SMART GMDH-NNM의 보편적인 해석과 최적 기상변수의 선정에 대한 결론을 유도하기 위해서는 지속적인 연구가 필요한 것으로 사료된다. 또한 증발접시 증발량을 비롯한 다양한 수문시계열의 모형화 예를 들면 강우 및 유출량의 모형화의 연구에도 가능한 것으로 판단되며, 이에 따른 지속적인 연구가 요구된다.

## 참고문헌

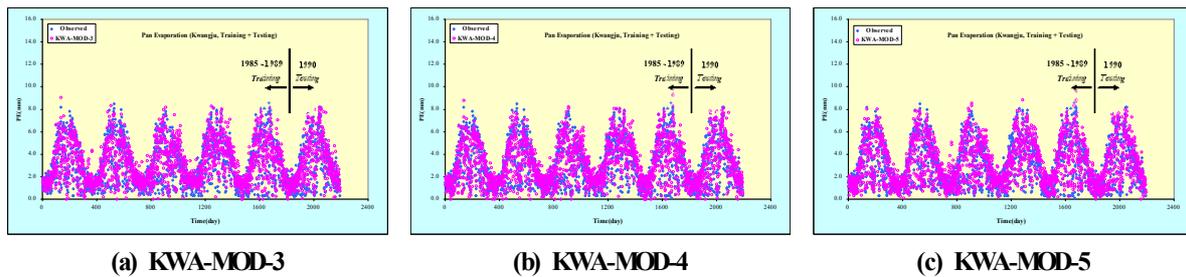
1. 국토해양부, 수자원 관리 종합정보 시스템 홈페이지 <http://www.wamis.go.kr>, 2009.
2. 기상청, 기상청 홈페이지 <http://www.kma.go.kr>, 2009.
3. Farlow, S.J. (1984). *Self-organizing method in modeling : GMDH type algorithms*, Statistics : Textbooks and Monographs, pp. 1-54.
4. Kim, S. and Kim, H.S. (2008a). "Uncertainty reduction of the flood stage forecasting using neural networks model." *J. Americ. Water Resour. Associ.*, Vol. 44, No. 1, pp. 148-165.
5. Kim, S. and Kim, H.S. (2008b). "Neural networks and genetic algorithm approach for nonlinear evaporation and evapotranspiration modeling." *J. Hydro.*, Vol. 351, pp. 299-317.
6. Kim, S., Kim, J.H., and Park, K.B. (2009). "Neural networks models for the flood forecasting and disaster prevention system in the small catchment." *Disas. Advn.*, Vol. 2, No. 3, pp. 51-63.
7. Lenters, J.D., Kratz, T.K., and Bowser, C.J. (2005). "Effects of climate variability on lake evaporation : Results from a long-term energy budget of Sparking lake, northern Wisconsin (USA)." *J. Hydro.*, Vol. 308, pp. 168-195.
8. Neuroshell 2. (1993). Ward systems group, Inc., MD.
9. Rahimi Khoob, A. (2009). "Estimating daily pan evaporation using artificial neural network in a semi-arid environment." *Theor. Appl. Climatol.*, Doi : 10.1007/s00704-008-0096-3.
10. Sabziparvar, A.A., Tabari, H., Aeni, A., and Ghafouri, M. (2009). "Evaluation of class A pan coefficient models for estimation of reference crop evapotranspiration in cold-semi arid and warm climates." *Water Resour. Manage.*, Doi : 10.1007/s11269-009-9478-2.
11. Terzi, O. and Keskin, M.E. (2005). "Modeling of daily pan evaporation." *J. Appl. Sci.*, Vol. 5, No. 2, pp. 368-372.
12. Vining, K.C. (2003). *Estimation of monthly evaporation from Lake Ashtabulain North Dakota, Orwell Lake in Minnesota, and Lake Traverse in Minnesota and South Dakota, 1931-2001*, Water Resources Investigation Report, U.S. Department of the Interior, U.S. Geological Survey.
13. Yoshimura, T., Deepak, D., and Tagaki, H. (1985). "Track/vehicle system identification by a revised group of data handling (GMDH)." *Int. J. Systems Sci.*, Vol. 16, No. 1, pp. 131-144.

**Table 1. SMART GMDH-NNM structure**

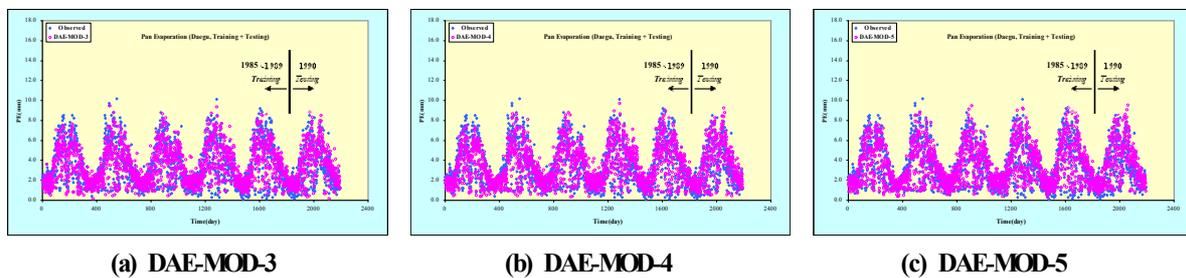
Station	Model	Input Nodes	Output Node
KWANGJU	KWA-MOD-1	$T_{max}$	PE
	KWA-MOD-2	$T_{max}$ , $T_{min}$	PE
	KWA-MOD-3	$T_{max}$ , $T_{min}$ , $W_{mean}$	PE
	KWA-MOD-4	$T_{max}$ , $T_{min}$ , $W_{mean}$ , $RH_{mean}$	PE
	KWA-MOD-5	$T_{max}$ , $T_{min}$ , $W_{mean}$ , $RH_{mean}$ , $SD$	PE
DAEGU	DAE-MOD-1	$T_{max}$	PE
	DAE-MOD-2	$T_{max}$ , $T_{min}$	PE
	DAE-MOD-3	$T_{max}$ , $T_{min}$ , $W_{mean}$	PE
	DAE-MOD-4	$T_{max}$ , $T_{min}$ , $W_{mean}$ , $RH_{mean}$	PE
	DAE-MOD-5	$T_{max}$ , $T_{min}$ , $W_{mean}$ , $RH_{mean}$ , $SD$	PE

**Table 2. Results of training performance**

Station	Model	Layers	Best Formula	Legend
KWANGJU	KWA-MOD-1	6	$Y=0.46*X1-0.61+0.16*X1^2-0.31*X1^3+0.18*X1^4+0.45*X1^5$	$X1=2.*(T_{max}+4.6)/40.6-1.$ $Y=2.*PE/12.5-1.$
	KWA-MOD-2	5	$Y=1.1*X1-0.69*X2-0.61+0.2*X1^2-1.3*X2^2+5.4E-002*X2^3+1.2*X1*X2+0.4*X2^4$	$X1=2.*(T_{max}+4.6)/40.6-1.$ $X2=2.*(T_{min}+12.5)/39.3-1.$ $Y=2.*PE/12.5-1.$
	KWA-MOD-3	9	$Y=1.7*X1-1.1*X2+0.11*X3-0.47+0.8*X1^2-0.8*X2^2-0.38*X3^2+7.6E-002*X2^3+0.23*X1*X2+0.86*X1*X3-0.67*X2*X3-0.89*X1*X2*X3+0.27*X1^3+0.27*X1*X2^2-0.29*X1*X3^2-6.7E-002*X1*X2^3-0.65*X1^2*X2+0.65*X1^2*X3-0.29*X1^2*X2*X3$	$X1=2.*(T_{max}+4.6)/40.6-1.$ $X2=2.*(T_{min}+12.5)/39.3-1.$ $X3=2.*(W_{mean}-2)/8.1-1.$ $Y=2.*PE/12.5-1.$
	KWA-MOD-4	12	$Y=-0.49*X2+0.14*X3-0.36-0.17*X4+1.2*X1+0.34*X1^2-0.72*X2^2-0.37*X4^2+4.2E-002*X2^3+0.78*X1*X2-0.29*X1*X4-0.16*X1*X3+0.45*X2*X3-0.21*X3^2-0.12*X1^2*X3-0.27*X2^2*X3+0.37*X3^3+5.E-002*X2^3*X3+0.75*X1*X2*X3-0.84*X1*X3^2+0.66*X2*X3^2+0.38*X1*X2*X3^2-0.21*X3*X4-0.49*X3*X4^2-0.34*X1*X3*X4-0.11*X4^3$	$X1=2.*(T_{max}+4.6)/40.6-1.$ $X2=2.*(T_{min}+12.5)/39.3-1.$ $X3=2.*(W_{mean}-2)/8.1-1.$ $X4=2.*(RH_{mean}-2.6)/93.$ $Y=2.*PE/12.5-1.9-1.$
	KWA-MOD-5	12	$Y=-0.2*X4+9.9E-002*X3+0.79*X1-0.13*X2-0.36+0.22*X5+0.25*X1^2-0.17*X4^2+0.21*X5^2-0.13*X1^3+9.5E-002*X4^3+0.23*X5^3+0.21*X1*X5+0.22*X1*X4*X5-8.E-002*X2^2+0.33*X3^3-0.11*X2*X3+0.29*X1*X3-8.6E-002*X3*X4+0.16*X3*X5+7.8E-002*X1^2*X3-0.11*X3*X4^2+0.12*X3*X5^2-9.6E-002*X1^3*X3+4.E-002*X3*X4^3+0.12*X3*X5^3+8.8E-002*X1*X3*X5+9.4E-002*X1*X3*X4*X5+7.6E-002*X4*X5-7.E-002*X2^3$	$X1=2.*(T_{max}+4.6)/40.6-1.$ $X2=2.*(T_{min}+12.5)/39.3-1.$ $X3=2.*(W_{mean}-2)/8.1-1.$ $X4=2.*(RH_{mean}-2.6)/93.9-1.$ $X5=2.*SD/15.77-1.$ $Y=2.*PE/12.5-1.$



**Figure 1. Comparison of pan evaporation for training and testing performance (Kwangju station)**



**Figure 2. Comparison of pan evaporation for training and testing performance (Daegu station)**